Tunneling problems by quantum Monte Carlo

Nikolai Prokof'ev, Boris Svistunov, and Igor Tupitsyn
Russian Research Center "Kurchatov Institute", 123182 Moscow, Russia

We develop a new numerical scheme which allows precise solution of coherent tunneling problems, i.e., problems with exponentially small transition amplitudes between quasidegenerate states. We explain how this method works for the single-particle (tunneling in the double-well potential) and many-body systems (e.g., vacuum-to-vacuum transitions), and gives directly the instanton shape and tunneling amplitude. Most importantly, transition amplitudes may be calculated to arbitrary accuracy (being limited solely by statistical errors) no matter how small are their absolute values.

73.40.Gk, 02.70.Lq, 05.30.-d

Tunneling phenomena are among the most intriguing consequences of quantum theory. They are of fundamental importance both for the high-energy and condensed matter physics, and the list of systems which behavior is governed by tunneling transitions ranges from quantum chromodynamics [1,2] to Josephson junctions and defects in crystals (see, e.g., Ref. [3]).

Precise analytic treatment of tunneling in complex systems is very hard, if not impossible. The most crucial simplification is in reducing the original problem to the semiclassical study of the effective action for some collective variable $\mathbf{R}$, with the assumption that all the other degrees of freedom adjust adiabatically to the motion of $\mathbf{R}$ [4,3]. In certain cases, one may also include dissipative effects due to "slow" modes other than the selected collective variable which *do not* follow the dynamics of $\mathbf{R}$ adiabatically [3,5]. Typically, the parameters of the effective action can not be found analytically (although one may relate some of them to the linear response coefficients) and have to be deduced from experiments. As far as we are aware, at present there are no tools to address the tunneling problem numerically. "Exact" diagonalization works only for relatively small systems, and, even in small systems, its accuracy is not sufficient to resolve very small energy splittings ΔE, say, when $\Delta E/E \ll 10^{-10}$, due to round-off errors (unless specific no-round-off arithmetics is used).

In this letter we develop a quantum Monte Carlo (MC) approach which allows precise calculations of tunneling amplitudes (and instanton shapes) no matter how small are their absolute values. The MC scheme contains no systematic errors, and its accuracy is limited only by statistical noise. The key point of our approach is in simulating imaginary-time dependence of transition amplitudes $A_{ji}(\tau)$ between selected reference states $|\eta_1\rangle$ and $|\eta_2\rangle$ (see below). In doing so we have to solve the problem of collecting reliable statistics in a case when $A_{ji}(\tau)$ varies, say, over hundreds of orders of magnitude (!) between different points in time. First, we describe in detail how to evaluate tunneling splitting and instanton shape in the double-well potential. We proceed then to the many-body problem of vacuum-to-vacuum transitions by considering the case of 1D quantum antiferromagnet with exchange anysotropy. Finally, we discuss the generality of the method suggested and demonstrate our numeric results for tunneling in the double-well potential (with comparison to the exact-diagonalization data where possible).

Consider the standard problem of particle motion in external potential:

$$H = m\dot{x}^2/2 + U(x) \; , \tag{1}$$

which has two minima at points $x = \eta_1$ and $x = \eta_2$, and large tunneling action $S = \int_{\eta_1}^{\eta_2} p\, dx = \int_{\eta_1}^{\eta_2} [2mU(x)]^{1/2} dx \gg 1$. These minima are supposed to be near-degenerate, i.e., the lowest eigenstates of the Hamiltonian (1), $H\Psi_\alpha = E_\alpha \Psi_\alpha$, form a doublet with

$$E_2 - E_1 \sim e^{-S}\omega_i \ll \omega_i \; , \tag{2}$$

where ω_i are the classical vibration frequencies in the potential minima $\omega_i = [U''(\eta_i)m]^{1/2}$ (we set $\hbar = 1$). The lattice analog of the Hamiltonian (1) reads

$$H = -t \sum_{<ll'>} d_l^\dagger d_{l'} + \sum_l n_l U_l \; , \quad n_l = d_l^\dagger d_l \; , \tag{3}$$

where $d_l^\dagger$ creates a particle on the site number l, and the first sum is over nearest-neighbor sites.

The transition amplitude from the state $|\eta_1\rangle$ to the state $|\eta_2\rangle$, where $|\eta_{1,2}\rangle = \delta(x - \eta_{1,2})$, in imaginary time τ is given by

$$A_{ji}(\tau) = \langle \eta_2 | e^{-H\tau} | \eta_1 \rangle \equiv \sum_\alpha \langle \alpha | \eta_1 \rangle \langle \eta_2 | \alpha \rangle e^{-E_\alpha \tau} \; . \tag{4}$$

We now make use of the inequality (2) to define the asymptotic regime $E_2 - E_1 \ll \tau^{-1} \ll \omega_i$, see, e.g., Refs. [6],

$$A_{ji}(\tau) \to e^{-\bar{E}\tau} \sum_{\alpha=1,2} \langle \alpha | \eta_1 \rangle \langle \eta_2 | \alpha \rangle [1 - (E_\alpha - \bar{E})\tau] \; . \tag{5}$$

where $\bar{E} = (E_2 + E_1)/2$.

It is convenient to split the double-well potential in two terms $U(x) = U^{(1)}(x) + U^{(2)}(x)$, where $U^{(1,2)}(x)$ is identical to $U(x)$ to the left/right of the barrier maximum point

and remains constant afterwards. Introducing system ground states in each minimum as $H^{(i)}\Psi_G^{(i)} = E_G^{(i)}\Psi_G^{(i)}$, where $H^{(i)} = m\dot{x}^2/2 + U^{(i)}(x)$, we may rewrite

$$\Psi_1 = u\Psi_G^{(1)} + v\Psi_G^{(2)} ,$$
$$\Psi_2 = v\Psi_G^{(1)} - u\Psi_G^{(2)} , \qquad (6)$$

where $(u^2, v^2) = 1/2 \pm \xi/2E$, $E^2 = \Delta^2 + \xi^2$, with obvious identification of the energy splitting $2E = E_2 - E_1$ and bias energy $\xi = (E_G^{(1)} - E_G^{(2)})/2$. Here Δ is the tunneling amplitude, which defines energy splitting $E_2 - E_1 = 2\Delta$ in the degenerate case. Substituting Eq. (6) into Eq. (5) we finally obtain

$$A_{ii}(\tau) \to e^{-\bar{E}\tau} Z_i^2 ,$$
$$A_{j\neq i}(\tau) \to \tau e^{-\bar{E}\tau} Z_i Z_j \Delta . \qquad (7)$$

Here $Z_i = \langle \Psi_G^{(i)} | \eta_i \rangle$ projects the reference states $|\eta_i\rangle$ on the corresponding ground states in each minimum. All corrections to Eq. (7), e.g., the neglect of the overlap integrals $\langle \Psi_G^{(1)} | \eta_2 \rangle$ and $\langle \Psi_G^{(2)} | \eta_1 \rangle$, are small in parameter e^{-S}. Note also that Eq. (7) does not depend on the bias $\xi \ll \omega_i$, and is insensitive to the behavior of $\Psi_G^{(i)}$ in the deep underbarrier region. This (rather standard) consideration relates transition amplitudes to the tunneling amplitude through the asymptotic analysis of $A_{ij}(\tau)$ in time.

MC simulation of the transition amplitude is almost identical to the standard simulation of quantum statistics [the partition function at a given temperature $T = 1/\beta$ may be written as $\mathcal{Z}(\tau = \beta) = \mathrm{Tr}_\eta A(\eta, \eta, \tau)$]. Fixed boundary conditions, as opposed to the trace over closed trajectories (configurations), are trivial to deal within any scheme. The real difference is in sampling different time-scales - thermodynamic calculations are typically done with $\beta = const$. Now we are forced to consider trajectories with different values of τ and to treat imaginary time as "dynamic" (in MC sense) variable. The idea of utilizing time dependencies of trajectories in MC simulations was extensively discussed in connection with the Worm algorithm [7,8] and polaron Green function [9].

We now turn to the problem of normalization. This problem might seem intractable in view of close analogy between $\mathcal{Z}(\tau)$ and $A(\tau)$. Formally, in the limit $\tau \to 0$, the amplitude is trivial to find in most cases, e.g., for the Hamiltonians (1) and (3) it is given by the free particle propagation

$$A_{ji}(\tau \to 0) \to \begin{cases} e^{-m(\eta_i - \eta_j)^2/2\tau} \sqrt{\frac{m}{2\pi\tau}} & \text{continuous} \\ (t\tau)^{|\eta_i - \eta_j|}/|\eta_i - \eta_j|! & \text{discrete} , \end{cases} \qquad (8)$$

and this knowledge may be used to normalize MC statistics for $A_{ji}(\tau)$ (in the discrete case $|\eta_i - \eta_j|$ =integer). However the absolute values of $A_{ji}(\tau)$ at short and long times will typically differ by orders and orders of magnitude and simply none MC statistics will be available at short times.

The solution to the puzzle lays in the possibility to use an arbitrary fictitious potential $A_{\mathrm{fic}}(\tau)$ in Metropolis-type updates [10] in time-domain

$$\frac{P_{acc}(\mathcal{B} \to \mathcal{A})}{P_{acc}(\mathcal{A} \to \mathcal{B})} = e^{-S_\mathcal{A} + S_\mathcal{B}} \frac{A_{\mathrm{fic}}(\tau')}{A_{\mathrm{fic}}(\tau)} \frac{W_\mathcal{A}(\tau)}{W_\mathcal{B}(\tau')} , \qquad (9)$$

where $P_{acc}(\mathcal{B} \to \mathcal{A})/P_{acc}(\mathcal{A} \to \mathcal{B})$ is the acceptance ratio for the update transforming initial trajectory $\mathcal{B}$ (having duration τ) to the trajectory $\mathcal{A}$ (having duration τ'), and $e^{-S_\mathcal{B}}$ is the statistical weight of the trajectory $\mathcal{B}$. The normalized distribution functions $W_\mathcal{B}(\tau')$, according to which a new value of τ is seeded, are also arbitrary; the best choice of W's follows from the conditions of (i) optimal acceptance ratio (as close to unity as possible), and (ii) simple analytic form allowing trivial solution of the equation $\int_{\tau_a}^{\tau'} W(\tau')d\tau' = r \int_{\tau_a}^{\tau_b} W(\tau')d\tau'$ on the time interval (τ_a, τ_b) , where $0 < r < 1$ is the random number [7,9]. Each trajectory adds a contribution $= 1/A_{\mathrm{fic}}(\tau)$ to the time-histogram of $A_{ij}(\tau)$.

One may use fictitious potential to enhance statistics of trajectories with certain values of τ "by hand", e.g., by making A_{fic} zero outside some time-window. To get a reliable and properly weighted statistics both at short and long times we need $A_{\mathrm{fic}}(\tau) \sim 1/A_{ij}(\tau)$ to compensate completely for the severe variation of the transition amplitude between different time-scales. This goal is achieved as follows. The initial stage of the calculation, called thermolization, prepares the fictitious potential using recursive self-adjusting scheme - starting from $A_{\mathrm{fic}}(\tau) = 1$ in a given time-window (τ_{min}, τ_{max}) and zero otherwise, we collect statistics for $A_{ij}(\tau)$ to the temporary time-histogram and after $M > 10^6 \div 10^7$ updates we renew the fictitious potential as

$$A_{\mathrm{fic}}(\tau) = \begin{cases} A_{ij}(\tau_0)/A_{ij}(\tau) , & \tau_1 < \tau < \tau_2 \\ A_{ij}(\tau_0)/A_{ij}(\tau_1) , & \tau_{min} < \tau < \tau_1 \\ A_{ij}(\tau_0)/A_{ij}(\tau_2) , & \tau_2 < \tau < \tau_{max} \end{cases} \qquad (10)$$

where τ_1 and τ_2 are the points (to the left and to the right of some reference point τ_0) where temporary statistics becomes unreliable and has large fluctuations. It makes sense to select τ_0 close to the maximum of $A_{ij}(\tau)$ (this point may be tuned a posteriory), and points τ_1 and τ_2 are formally defined as the first points in the histogram where smooth variation of $A_{ij}(\tau)$ ends: $A_{ij}(\tau_{1,2} + \Delta\tau)/A_{ij}(\tau_{1,2}) < \delta$ (here $\Delta\tau$ is the difference between the nearest points in time-histogram, and δ is a small number, say 0.01). The thermolization stage continues until $\tau_1 = \tau_{min}$ and $\tau_2 = \tau_{max}$, and fictitious potential stops changing (withing a factor of two). After that the actual calculation starts with a fixed $A_{\mathrm{fic}}(\tau)$, and a new histogram for $A_{ij}(\tau)$ is collected.

The idea of using fictitious potential proportional to the inverse of the transition amplitude is clear - it allows to collect reliable statistics on different time-scales with comparable relative accuracy. With this tool at hand one

can easily normalize $A_{ij}(\tau)$ using known analytic results for short times, and deduce tunneling amplitude and Z-factors from the analysis of the long-time asymptotic, Eq. (7). We can not but note that fictitious potential in time-domain very much resembles the so-called "guiding wavefunction" in the Green-function MC methods [11], with essential difference that here it is used to reach exponentially rare configurations.

Instanton shape calculation is a much easier task since it can be done by considering trajectories for the transition amplitude $A_{i\neq j}(\tau)$ with fixed but sufficiently long $\tau \gg 1/\omega_i$; again, to ensure that A_{ij} is dominated by just one instanton trajectory we need $\tau \ll 1/\Delta$. For any given MC trajectory $x(\tau)$ one has first to define the instanton center position in time, and to recount all times from this center-point τ_c [instanton center statistics is almost uniform in $(0,\tau)$ (except near the ends of the time interval) due to the generic "zero-mode" present in instanton solutions [2,6]]. This can be done by looking at the average time

$$\tau_c = \frac{\int_B \tau d\tau}{\int_B d\tau} \ , \tag{11}$$

where the integral is taken over the barrier region between the wells $U(x(\tau)) > E_G$. The instanton shape is obtained then by collecting statistics of $x(\tau - \tau_c)$ to the time histogram. In this simple example when the notion of collective coordinate is not necessary (or formally, $\mathbf{R} = x$), we do not need to define separately the estimator for $\mathbf{R}$ (see the opposite example below).

The case of vacuum-to-vacuum transition in the many-body problem is formulated in precisely the same manner, and Eqs. (4-7), (9-10) holds true once the identification of the states $|\eta_{1,2}\rangle$ is done and the short-time limit, Eq. (8), is calculated. Consider as a typical example a 1D spin chain with $2L$ sites and antiferromagnetic (AF) couplings between the nearest-neighbor spins

$$H = \sum_{<ij>} \left[J\vec{S}_i\vec{S}_j + J' S_i^z S_j^z \right] \ , \tag{12}$$

which has near degenerate ground state with the lowest doublet separated from the rest of the system spectrum by finite gap for $J' > 0$. The natural choice of $|\eta_{1,2}\rangle$ is then an ordered AF state with $S_i^z|\eta_{1,2}\rangle = (\pm 1)^i|\eta_{1,2}\rangle$. Note, that for a large system Z-factors $\langle \Psi_G^{(i)}|\eta_i\rangle$ will be also exponentially small, and even diagonal amplitudes A_{ii} have to be calculated with the use of $A_{\rm fic}$ [in the single-particle case one may obtain Z-factors by ignoring $A_{\rm fic}$-trick]. The short-time behavior is given by

$$\begin{aligned} A_{ii}(\tau) &\to 1 \\ A_{i\neq j}(\tau) &\to (J\tau/2)^L \begin{cases} 2 & \text{(ring)} \\ 1 & \text{(open chain)} \ , \end{cases} \end{aligned} \tag{13}$$

To decipher the instanton we need now some *a priori* knowledge about the relevant collective variable (if such knowledge is not available one has to study different possibilities). For example, if tunneling proceeds via two domain walls well-separated from each other (thin-wall approximation [4,2,6]), then the collective variable $\mathbf{R}$ is the distance between the walls, and the underbarrier region in Eq. (11) is related to the existence of two separated walls. These definitions are not too specific and work only approximately; this is however the generic difficulty of dealing with collective variables which are meaningful only in the macroscopic limit. Obviously, the knowledge of the instanton shape does not allow precise evaluation of Δ, and gives only rough estimation of $\ln \Delta$.

To test the proposed scheme and to compare results to the exact diagonalization (ED) data we have applied our algorithm to the lattice model (3) with $U(x) = U_0[(x/\eta)^2 - 1]^2$. In what follows we measure all energies in units of the hopping amplitude t and count them from the potential minimum. We set $U_0 = 1$ and consider two interwell separations: $\eta = 10$ and $\eta = 40$. For $\eta = 10$ the ED data for the ground state energy, Z-factor, and tunneling amplitude are: $E_G = (E_1 + E_2)/2 = 0.1923$, $Z^2 = 0.2465$, $\Delta = (E_2 - E_1)/2 = 3.6078 \times 10^{-6}$. Our MC data give $E_G = 0.192(2)$, $Z^2 = 0.246(2)$, and $\Delta = 3.61(3) \times 10^{-6}$. The case of large η is more subtle since only $E_G = 0.0495$ and $Z^2 = 0.1255$ may be tested against ED - one may use textbook semiclassical analysis of the corresponding continuous model to see that $\Delta(\eta = 40) \sim 10^{-24} \div 10^{-23}$, that is far beyond the standard computer round-off errors.

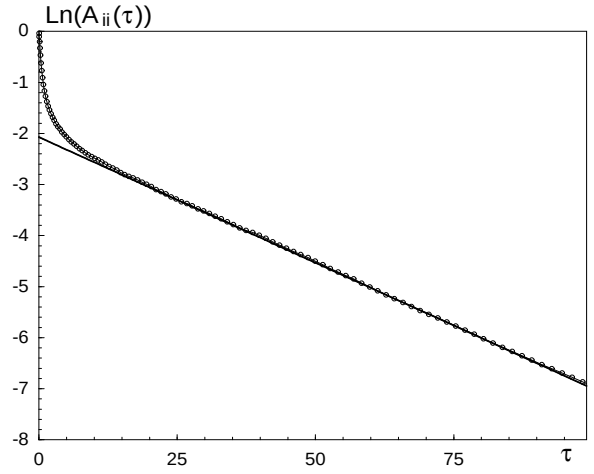

FIG. 1. Time dependence of the diagonal amplitude A_{ii}. Solid line $Z^2 e^{-E_G \tau}$ is the prediction of the long-time behavior according to the ED data.

In Fig. (1) and Fig. (2) we present our MC data for the diagonal and off-diagonal amplitudes for $\eta = 40$, and fits to the expected long-time, Eq. (7), and short-time behavior, Eq. (8) [we also include the lowest-order correction for the potential energy at $\tau \to 0$ which tells that $A_{i\neq j} \propto e^{-\bar{U}\tau}$ where $\bar{U} = \int_{-\eta}^{\eta} U(x)dx$]. The variation

of $A_{i\neq j}$ in Fig. 2 is about two-hundred orders! From Fig. (1) we deduced $E_G = 0.0494(2)$ and $Z^2 = 0.126(1)$ in agreement with ED. From Fig. (2) we then obtain $\Delta = 1.7(4) \times 10^{-23}$. MC simulation of the above figures took about 5 days each on PII-266 processor.

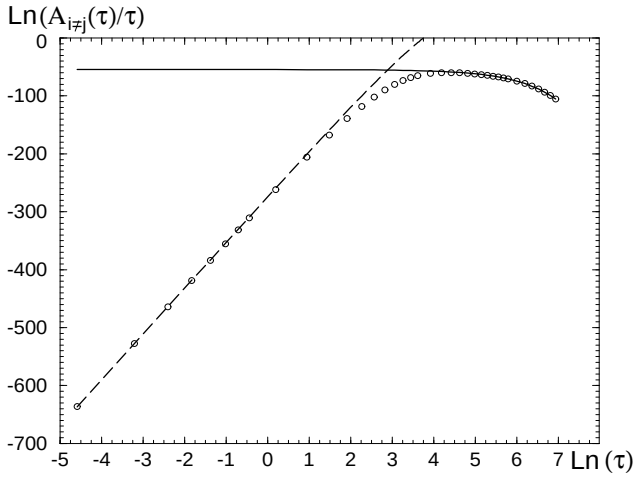

FIG. 2. Time dependence of the non-diagonal normalized amplitude $A_{i\neq j}$. The solid line is fitting to the curve $\Delta Z^2 e^{-E_G\,\tau}$ and the short-time behavior is tested against the law $(\tau^{2\eta-1}/(2\eta)!)e^{-\bar{U}\tau}$ (dashed curve).

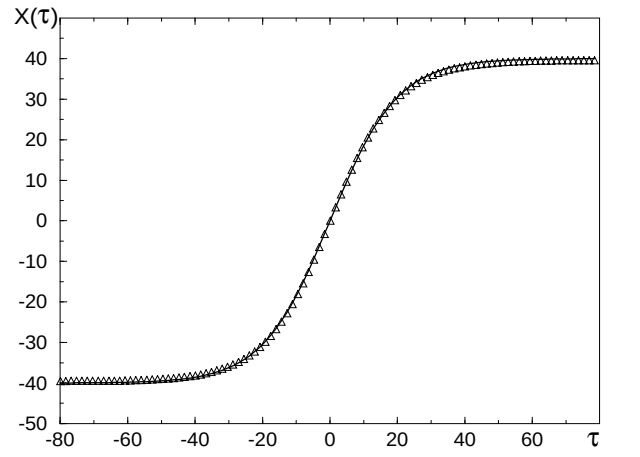

FIG. 3. Instanton trajectory $x(\tau)$. The solid curve $x(\tau) = \eta \tanh(E_G\,\tau)$ is the semiclassical result for the continuous model.

In Fig. (3) we present our results for the instanton trajectory $x(\tau)$ (at $\eta = 40$) along with the known semiclassical results for the continuous model (1) [2]. The accuracy of the data is self-evident, although we argue that MC data rather represent the trajectory with finite energy $E = E_G$ while analytic results correspond to $E = 0$.

We note that the present technique is very hard to implement in the discrete-time schemes with finite Trotter parameter $\Delta\tau$. On one hand, long-time asymptotic regime (see Fig. 2) requires to consider τ as long as 640. On another hand at short times, the requirement of smooth variation of the amplitude $\Delta\tau\, d\ln[A_{i\neq j}(\tau)]/d\tau \ll 1$, after substituting Eq. (8), means $\Delta\tau \ll \tau/(2\eta) \approx 10^{-3}$ for $\tau = 0.1$. To avoid large systematic errors due to time-discretization at short times one has to use $\Delta\tau = 10^{-4}$(!!). Apart from enormous memory usage (there will be about 10^7 time slices) that small Trotter parameter severely slows down the efficiency of the code, in fact, we are not aware of any MC simulation with $\Delta\tau \sim 10^{-4}$.

It is worth mentioning that similar technique makes it possible to study directly $\mathcal{Z}(\tau)$ over different time-scales - normalization of the partition function does not matter since none physical quantity depend on it. Thus one can obtain temperature dependences of the free energy, entropy etc. in a single MC run.

This work was supported by the RFBR Grants 98-02-16262 and 97-02-16548 (Russia), IR-97-2124 (European Community).

[1] A.A. Belavin, A.M. Polyakov, A.S. Schwartz, Yu.S. Tyupkin, Phys. Lett. **B**, **59**, 85 (1975).

[2] A.I. Vainshtein, V.I. Zakharov, V.A. Novikov, and M.A. Shifman, Usp. Fiz. Nauk., **136**, 553 (1982); M.A. Shifman, *Instantons in Gauge Theories*, World Scientific, Singapore (1994); R. Rajaraman, *Solitons and Instantons*, North-Holland, Amsterdam (1987).

[3] *Quantum Tunneling in Condensed Matter*, eds. Yu. Kagan and A.J. Leggett, North-Holland, Elsevier (1992).

[4] I.M. Lifshitz and Yu.M. Kagan, Sov. Phys.-JETP, **35**, 206 (1972); S.V. Iordanski and A.M. Finkelshtein, Sov. Phys.-JETP, **35**, 216 (1972); B.V. Petukhov and V.L. Pokrovskii, Sov. Phys.-JETP, **36**, 340 (1972).

[5] U. Weiss, *Quantum Dissipative Systems*, Ser. in Modern Condensed Matter, World Scientific, Singapore, (1998).

[6] S. Coleman, *The Uses of Instantons*, in *The Whys of Subnuclear Physics*, ed. A. Zichichi, Plenum Press, New-York (1979).

[7] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Zh. Eksp. Teor. Fiz., **114**, 570 (1998).

[8] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Phys. Lett. **A**, **238**, 253 (1998).

[9] N.V. Prokof'ev and B.V. Svistunov, Phys. Rev. Lett. **81**, 2514 (1998).

[10] N. Metropolis *et al.*, J. Chem. Phys., **21**, 1087 (1953).

[11] D.M. Ceperley, J. Comp. Phys., **51**, 404 (1983); D.M. Ceperley and B.J. Alder, J. Chem. Phys., **81**, 5833 (1984).